

Sequence Listing

Individual Applicant

Street :
City : Athens
State : GA
Country : USA
PostalCode : 30602-7272
PhoneNumber : 706-542-0928
FaxNumber : ____-_____
EmailAddress : wparrott@uga.edu
<110> LastName : Parrott
<110> FirstName : Wayne
<110> MiddleInitial : A
<110> Suffix : PhD.

Individual Applicant

Street :
City : Athens
State : GA
Country : USA
PostalCode : 30602-7272
PhoneNumber : 706-542-0928
FaxNumber : ____-_____
EmailAddress :
<110> LastName : LaFayette
<110> FirstName : Peter
<110> MiddleInitial : R
<110> Suffix : PhD.

Individual Applicant

Street : 1907 South Milledge Ave
City : Athens
State : GA
Country : USA
PostalCode : 30605-_____
PhoneNumber : ____-_____
FaxNumber : ____-_____
EmailAddress :
<110> LastName : Kane
<110> FirstName : Patrick
<110> MiddleInitial : M
<110> Suffix :

Application Project

<120> Title : Arabitol or Ribitol As Positive Selectable Markers
<130> AppFileReference : UGA-855R
<140> CurrentAppNumber :
<141> CurrentFilingDate : ____-____-

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :
atgaacgaac aatttacatg gctgcacatc gggtaggtt ctttcatcg cgcacatcag 60
gcgtggatc tacaccgtt gcaggtatg ggcataaac gctggagcat tgctgoggc 120
aatattcgtt atgatgctga acatgtcgta caggcactca gtgcacagaa aggtcgat 180
gtgctggaaa ccgtcagccc ggaaggggta agcgaatatg aagagatcac ctcaattcag 240

aagttgatac	cgtggcaggc	agatttacaa	ccgctgattg	ctgaaggggc	agatccgaag	300
acaaaagtga	ttgctttcac	cgtcaccgaa	ggcgggtact	acctgaatac	cagtcacaaa	360
ctggaaagtt	acaatcctga	tttagcggca	gatcttaaag	ggggatgcaa	aacaatttac	420
ggtgttatta	cccgatccct	cgaagcgcgt	atggcaaata	acgcccggacc	actaaccctg	480
atgaacgaac	aatttacatg	gctgcacatc	gggttaggtt	ctttcatcg	cgcacatcg	540
gcgtggtatac	tacaccgtt	gcaggtgatg	ggcgataaac	gctggagcat	tgctgcgggc	600
aatattcgt	atgatgctga	acatgtcgt	caggcaactca	gtgcacagaa	aggtcgtat	660
gtgctggaaa	ccgtcagccc	ggaaggggtt	agcgaatatg	aagagatcac	ctcaatttcg	720
aagttgatac	cgtggcaggc	agatttacaa	ccgctgattt	ctgaaggggc	agatccgaag	780
acaaaagtga	ttgctttcac	cgtcaccgaa	ggcgggtact	acctgaatac	cagtcacaaa	840
ctggaaagtt	acaatcctga	tttagcggca	gatcttaaag	ggggatgcaa	aacaatttac	900
ggtgttatta	cccgatccct	cgaagcgcgt	atggcaaata	acgcccggacc	actaaccctg	960
ctgaatttgcg	ataacgtgcg	ccataatgtt	gaacgttcc	atgatggctt	ggttgagttt	1020
ctccagctaa	ctggcaaaca	ggatgtcata	gactggctg	gtacaaatac	cacttgcggc	1080
aataccatgg	ttgaccgcata	tacgcctgt	ccggcagcag	aaatccggc	acggatcaag	1140
gctcaaacgg	gtattgccc	taaagcgcg	gtaatggcg	aaacctttat	ccagtgggtc	1200
gtggaagata	atttccgtt	tgtccgtcc	gcactggaga	aggtcgggtt	cgaactgggt	1260
gcgtcggtaa	tcccctatg	agaggcgaag	attcgcattt	ttaactctt	acacaggttc	1320
atcgcttggg	caggatcata	aatcggttca	aaatataatcc	acgaaagcac	aatgaccgt	1380
tttatctatc	agatttgcga	ccgctacgt	acggaagatg	tcatttcctt	cttggggcgat	1440
aacggatatcg	attttgcac	ctaccgtat	gttgtactca	acgcgttttac	caatccacat	1500
attcaggaca	ccaaccaacg	cgtcgctgcg	gatgggttct	cggaaattcc	ggcgatgatt	1560
gccccccacac	tgcgagagtg	ctaccagcga	ggcggtcgcc	cgaatgccac	cggccatgtt	1620
cctgcactgt	tttacgtatt	catggagcag	tggcatcact	gcaaaactgcc	ctatgaatat	1680
caggatggca	tccttgcgtt	accagctgtc	catgcataat	tacagtcgtc	cgatcccgcc	1740
gctgtttatg	ccagtgataa	agcgctgtt	ggcgattta	ccgaacgtga	agattttgcc	1800
gcgttgcgtt	gcgaaaaaaat	cgctgacgtc	tacgcgttta	ttaacttaa		1848

<212> Type : DNA

<211> Length : 1848

SequenceName : SEQ ID NO 1

SequenceDescription : arabitol dehydrogenase

Custom Codon

Sequence Name : SEQ ID NO 1

Sequence

<213> OrganismName : Escherichia coli

<400> PreSequenceString :

atcgatttgcgtt	cacacggcag	ctaaattccc	gttcagtgcg	tgcaaggaaa	60
cagagactat	aaattcgcggc	ttgtaaaagg	attatatgt	gaatcaactt	120
tgaataactcc	ccttaatggc	aaagggtcag	ctatcaactgg	cgtcgctgc	180
tgcaatgtgc	aaaaaacgtt	ctcgatgcag	gagcaaaagg	atgatttattt	240
gcgacaaact	gcacaagatt	gtcgctgtt	taggcggaaa	cgcgtacgc	300
atcttccaa	taatcagcaa	gtcgatcaaa	tgctggcgga	cattatcgaa	360
ggctggatat	ttttcatgca	aatgcaggcg	cttataatgg	cggcccagg	420
atccagatgt	ctggatgtt	gtgtttaatc	tgaatataaa	tgccgcgtt	480
gtgcagtcct	gcccatatg	attgcgcaga	ggtcggggcg	tataattttt	540
tcgcggggcg	cggtccgggtt	atctggaaac	cgatcttacac	tttgcgttcc	600
aggcatttgt	acacactacc	cgccgcagg	tttctcaata	tggcggtcg	660
tgctgcagg	accagtagtc	actgcctgc	ttgatgactt	gccaaggccc	720
aaggccctggc	aaatggtagc	ctgatgcac	cgatggaaat	ggcggatca	780
tggtgacccg	ctcgaaaaaaat	gtcaccgtc	gagatttagt	gatcctgcct	840
atctgttgc	gcgcaatcat	gacaataacc	aaaaccgtt	tttgcgttga	900
ggcagtgtcc	gcgcggggat	ttttgtatc	aacggatctc	tgcataatccc	960
aaaatcacga	ctacgcggcg	cagcggaaac	cgcggttgc	ggagatctgg	1020
caggcggctt	gttcatgtat	tcgtatgtc	ctcaactctgg	cagacgtttt	1080
gtggcaggca	tcgggtttga	tgccacatgt	tctctgttgg	tactggataa	1140
ccattgcctg	tcagccccgg	aggagatgc	aagcaaaata	tcatgtgt	1200
cgcgcaccgg	aacaaggcaga	gcgaatcaat	gccactcacc	atccgggtct	1260
ggtggtaaaa	tttgcgttca	aatggaaaca	ccgaaaattt	tctggctgaa	1320
ccagagatct	acgaacgtgc	cggacaattt	ttcgatctgg	ccgatttct	1380
gctaccgggt	atttagcgcg	ttcgtatgtc	actgttaccc	gacctggcc	1440

catgaaaatc	gctgggatcc	agattatttc	cgcäccattg	gccttgcaga	gttagcggat	1500
gaagattta	ttcgcattgg	tcatcatatt	gtttctcccg	gaacacccctg	tggaaatggt	1560
ttaacagcac	aaggccggc	agagatggg	ttactcccg	gcacacccgt	cgctgttaggg	1620
ttaattgatg	ctcacgctgg	tggcatcggt	acggttggc	tagaagggtgg	agcgtctgaac	1680
aatctcgct	atgtttcgg	cacttctca	tgcaccatgg	catctaccac	ttctccctcg	1740
tttgcacgg	gtgtctgggg	gccgttattac	agtgcgttgg	ttccaggtct	gtggtttagtt	1800
gaaggcgggc	aaagtgtgc	gggagcagct	attgaccagc	tacttgattt	ccatccggct	1860
gttgaagaag	ctcgcaaat	ggcacaacgt	gtgaatcagc	ccctcccccgt	ctggcttgc	1920
gatcgaatcc	tcgaaaaaaac	ggcgcaacca	tcagatgctg	tcgcccggc	gaaagggtca	1980
cacgtggtgc	cggaaattct	cggaaatcgc	gcccccttcg	cagatcctca	tgccagagcg	2040
gtaatttgc	gcctgggtat	ggagcgttat	ctggataatt	tactcgccctt	gtatatcgct	2100
ggattatgcg	gaatgggtt	tggctctgcgc	caaattctcg	acgtcaaaac	agcgcaggga	2160
gtatgtgat	aaaatatcg	tattagcggc	ggtgccgggc	agcatccact	gttacgacaa	2220
attctggcag	atacctgcgg	tattccggt	attaccacgc	aatgctgcga	accggttta	2280
ttaggctcgg	ccattcttgg	tgctgtcgcc	ggaaatatttgc	cacccctgtt	tggcgaagcg	2340
atgcaacaat	tcacccatgt	ggataaatat	tattatccgc	aagaacgcata	tcaatctt	2400
catcatcg	gatatgaggc	ctataaggcag	ttgcagcata	ctgcacaaatt	actcagagac	2460
taattaacca	gcccctgc	gctgtttca	ggcaatcaact	aataacgact	cactccggta	2520
atatcccgga	gtgcattcat	ctgcacccctt	aaaacgagggt	ttatatgtcc	agaaataata	2580
aacagtgggtt	gggtttgcca	ctgcatctga	tatggggata	tatcgccatc	gcagtattt	2640
tgactgggtt	tgggttcgaa	ctcgatttct	tatcgacta	tattaaatcg	ttaggcttca	2700
cacccgcgg	agccctttt	gcctttacgc	tctacggct	ggcggctgccc	ctttccgcct	2760
gggtttctgg	ggtagtagcg	gaaatcatca	cgccgcacaaa	aaccatgttgc	attggtttttgc	2820
tcctatgggt	cgtattccat	gttctgttgc	tggcttttgc	attaggacag	gcaaaactatg	2880
ggttaatct	gctgttttac	gggatttgcgt	gttttagcgt	tccgctattt	ctctactcat	2940
ttatcgctgt	tattattcat	aacgtgcgt	gcgaaaatttgc	cagttctgtt	ctgggtttgg	3000
actggggcgtt	atattcgttgc	gggatcggttgc	tggctggcag	ttatattccc	agttttacgtt	3060
taccgattat	gggtgaaatg	ggAACCTTAT	ggctggcaact	ggcgttctgc	tttgcggcgt	3120
gtgtcattgc	catgatctcc	ttgcgtcatg	ttaaaaacgc	tggacatatg	cataattaa	3180
ctccccgtga	gaagtttgca	gaattaagtc	gggcagtaac	tttactttat	accaaccgca	3240
atattttct	ctccaggattt	gtgcgcattt	taaataccctt	atcgatttttgc	gtttttgcgg	3300
tcattatgcc	aatgatgttt	gtggatgaac	tggatttgcac	cacccctgttgc	tggttgcagg	3360
tctggccggc	atttttctt	accactattt	tctctaaat	ttttttgggg	attgtggcag	3420
aaaaaaatggg	ctggatgcgt	gttatttgcgt	ggttttttgc	cctcggggat	gcagcatcaa	3480
gttttagcgtt	ttactacatgt	ccgcaataact	ttggtcacaa	ctactggat	gcaatgatttgc	3540
cggcgatttgc	tctgggaact	tttgcgttgc	catttgcgc	gatggccgt	gtcttcccg	3600
cactggaaacc	aaaacacacaa	ggtgcgttgc	tctcggttta	caaccccttct	gcgggtatgt	3660
ctaacttcc	ggctccggca	attgcgttgc	tgttattacc	gtgggtttagt	actatcggt	3720
tggtcattgc	ctatacagca	ttgttatctat	ttggcccttgc	ccttgcgc	ttcattcg	3780
ttgagcagcc	aggatttgcgt	tctgcgttgc	tgacttgagaa	ggcatttgc	atctccgttgc	3840
aaacgaaacg	catcaggcac	tcatcccttc	cctcatggg	gaggatgatt	tcacatcagg	3900
caatagtgc	tttgcgttgc	agataaaacgt	cctgcacggc	gttaatcgt	ttcacccgt	3960
cagccatcgat	t					3971

<212> Type : DNA

<211> Length : 3971

SequenceName : SEQ ID NO 2

SequenceDescription : ribitol operon

Feature

Sequence: SEQ ID NO 2:

<221> FeatureKey : misc_feature

<222> LocationFrom : 96

<222> LocationTo : 848

Other Information : ribitol dehydrogenase coding region

CDSJoin : No

Feature

Sequence: SEQ ID NO 2:

<221> FeatureKey : misc_feature

<222> LocationFrom : 859

<222> LocationTo : 2463

Other Information : ribitol kinase coding region

CDSJoin : No

Feature

Sequence: SEQ ID NO 2:

<221> FeatureKey : misc_feature
<222> LocationFrom : 2565
<222> LocationTo : 3839

Other Information : ribitol transporter coding region

CDSJoin : No

Custom Codon

Sequence Name : ribitol operon

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :

MMNHSVPSMN	TPLNGKVAAL	TGAASGIGLQ	CAKTLIDAGA	KVVLIDREGD	KLHKIVAEGL	60
ENAYALQLDL	FNNQQVNDNML	ADIIELAGGL	DIFHANAGAY	IGGPVAEGDP	DVWDRVLNLN	120
INAAFRCVRA	VLPHMTAQRS	GDIIFTSSIA	GVVPIWEPI	YTASKFAVQA	FVHTTRROVS	180
QYGVVRGAVL	PGPVVTALLD	DWPKAQMEAA	LANGSLMQPI	EVAESVLFMV	TRSKNVTVRD	240
LVLPGSVDL						250

<212> Type : PRT

<211> Length : 250

SequenceName : SEQ ID NO 3

SequenceDescription : ribitol dehydrogenase

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :

MTITKTVIGV	DVGSGSVRAG	IFDLNGSLLS	HATEKITTTR	RSGSRVEQSS	QEIWQAVCSC	60
IRNALTLADV	CAQSVAIGF	DATCSLVVLD	KNGDPLPVSP	EGDAKQNIIV	WMDHRATEQA	120
ERINATHHPV	LNYVGGKISP	EMETPKILWL	KENMPPEIYER	AGQFFDLADF	LTWRATGDLA	180
RSVCTVTCKW	TWLAHENRWD	PDYFRTIGLA	ELADEDFIRI	GHHVSPGTP	CGNGLTAQAA	240
AEMGLLPGTP	VAVGLIDAHA	GGIGFTVGVEG	GALNNLAYVF	GTSSCTMAST	TSPSFVPGVW	300
GPYYSAMVPG	LWLVEGGOSA	AGAAIDQLLD	FHPAVEEARE	MAQRVNQPLP	VWLADRILEK	360
TAQPSDAVAL	AKGLHVVPF	LGNRAPFADP	HARAVICGLG	MERDLDNLLA	LYIAGLCGIG	420
YGLRQILDAQ	TAQGVVSKNI	VISGGAGQHP	LVRQILADTC	GIPVITTQCC	EPVLLGSAIL	480
GAVAGNIAPS	VGEAMQQFTH	VDKYYYPQER	YQSLHHRRYE	AYKQLQHTAK	LLRD	534

<212> Type : PRT

<211> Length : 534

SequenceName : SEQ ID NO 4

SequenceDescription : ribitol kinase

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :

MSRNNKQWLG	LPLHLLWGYI	AIAVFMTGDG	FELAFLSHYI	KSLGFTPAEA	SFAFTLYGLA	60
AALSAWVSGV	VAEIITPQKT	MLIGFVLWCV	FHVLFLVFLG	GQANYGLILL	FYGIRGLAYP	120
LFLYSFIVVI	IHNVRSENSS	SALGWYWAVY	SVGIGVAGSY	IPSFTIPIMG	EMGTLWLALA	180
FCFAGGVIAM	ISLRHVKTPG	HMHNLTREK	FAELSRAVTL	LYTNRNIFLS	SIVRIINTLS	240
LFGFAVIMPM	MFVDELGFTT	SEWLQVWAAT	FFTTIFSNIF	WGIVAEKMGW	MRVIRWFGL	300
GMAASSLAFY	YMPQYFGHNY	WMAMIPAIAL	GTFVAAFPVM	AAVFPALPK	HKGAAISVYN	360
LSAGMSNFLA	PAIAVVLLPW	FSTIGVVIAY	TALYLLAFVL	CAFIRVEQPG	FSSAPVTEKA	420
LNIS						424

<212> Type : PRT

<211> Length : 424

SequenceName : SEQ ID NO 5

SequenceDescription : ribitol transporter